Published in final edited form as:
Ear Hear. 2009 August ; 30(4): 411–418. doi:10.1097/AUD.0b013e3181a61bc0.

Development and Validation of the University of Washington Clinical Assessment of Music Perception Test

Robert Kang[1,2], **Grace Liu Nimmons**[3], **Ward Drennan**[2], **Jeff Longnion**[2], **Chad Ruffin**[2,4], **Kaibao Nie**[1,2], **Jong Ho Won**[2,5], **Tina Worman**[1], **Bevan Yueh**[6], and **Jay Rubinstein**[1,2,5]

[1]Department of Otolaryngology–Head and Neck Surgery, University of Washington, Seattle, Washington

[2]Virginia Merrill Bloedel Hearing Research Center, University of Washington, Seattle, Washington

[3]Department of Otolaryngology–Head and Neck Surgery, University of Iowa, Iowa

[4]Louisiana State University School of Medicine, Shreveport, Louisiana

[5]Department of Bioengineering, University of Washington, Seattle, Washington

[6]Department of Otolaryngology–Head and Neck Surgery, University of Minneapolis, Minnesota

Abstract

Objectives—Assessment of cochlear implant outcomes centers around speech discrimination. Despite dramatic improvements in speech perception, music perception remains a challenge for most cochlear implant users. No standardized test exists to quantify music perception in a clinically practical manner. This study presents the University of Washington Clinical Assessment of Music Perception (CAMP) test as a reliable and valid music perception test for English-speaking, adult cochlear implant users.

Design—Forty-two cochlear implant subjects were recruited from the University of Washington Medical Center cochlear implant program and referred by two implant manufacturers. Ten normal-hearing volunteers were drawn from the University of Washington Medical Center and associated campuses. A computer-driven, self-administered test was developed to examine three specific aspects of music perception: pitch direction discrimination, melody recognition, and timbre recognition. The pitch subtest used an adaptive procedure to determine just-noticeable differences for complex tone pitch direction discrimination within the range of 1 to 12 semitones. The melody and timbre subtests assessed recognition of 12 commonly known melodies played with complex tones in an isochronous manner and eight musical instruments playing an identical five-note sequence, respectively. Testing was repeated for cochlear implant subjects to evaluate test-retest reliability. Normal-hearing volunteers were also tested to demonstrate differences in performance in the two populations.

Results—For cochlear implant subjects, pitch direction discrimination just-noticeable differences ranged from 1 to 8.0 semitones (Mean = 3.0, SD = 2.3). Melody and timbre recognition ranged from 0 to 94.4% correct (mean = 25.1, SD = 22.2) and 20.8 to 87.5% (mean = 45.3, SD = 16.2), respectively. Each subtest significantly correlated at least moderately with both Consonant-Nucleus-Consonant (CNC) word recognition scores and spondee recognition thresholds in steady state noise and two-talker babble. Intraclass coefficients demonstrating test-retest correlations for pitch, melody, and timbre were 0.85, 0.92, and 0.69, respectively. Normal-hearing volunteers had

---

Address for correspondence: Jay T. Rubenstein, MD, PhD, Virginia Merrill Bloedel Hearing Research Center, University of Washington, Box 357923 Seattle, WA 98195-7923. rubinj@u.washington.edu.

a mean pitch direction discrimination threshold of 1.0 semitone, the smallest interval tested, and mean melody and timbre recognition scores of 87.5 and 94.2%, respectively.

Conclusions—The CAMP test discriminates a wide range of music perceptual ability in cochlear implant users. Moderate correlations were seen between music test results and both Consonant-Nucleus-Consonant word recognition scores and spondee recognition thresholds in background noise. Test-retest reliability was moderate to strong. The CAMP test provides a reliable and valid metric for a clinically practical, standardized evaluation of music perception in adult cochlear implant users.

INTRODUCTION

The assessment of cochlear implant outcomes revolves around speech discrimination. Cochlear implants are designed to transmit aspects of the sound signal considered to be characteristic of speech, and various coding strategies exist that attempt to transmit speech effectively. Therefore, audiologic outcomes after cochlear implantation have been measured primarily with word and sentence recognition tests such as the Consonant-Nucleus-Consonant (CNC; Peterson & Lehiste 1962) and Hearing in Noise test (Nilsson et al. 1994; Luxford 2001). Cochlear implant users have demonstrated remarkable speech perception ability, because the majority of postlingually deafened implantees using current devices achieve close to 80% on sentence recognition tests in quiet listening conditions (Wilson 2000; Friesen et al. 2001). Despite these significant improvements in speech perception, basic elements of music perception remain a challenge for most cochlear implant users (Gfeller & Lansing 1991; Leal et al. 2003; Kong et al. 2004). Many implant users have expressed disappointment and frustration at how poorly music sounds through their implants and desire to better appreciate music (Gfeller et al. 1998, 2000).

Music can be dissected into several fundamental components such as rhythm, pitch, melody, and timbre. The ability to perceive each of these basic elements of music has been tested in cochlear implant users. Implant users have demonstrated the ability to discriminate simple rhythmic patterns much more easily than pitch or melodic patterns, even at levels similar to normal-hearing listeners (Gfeller & Lansing 1991; Gfeller et al. 1997). Pitch discrimination has proved to be a more challenging task for cochlear implant users (Gfeller et al. 1997, 1998, 2000). Gfeller et al. found that the ability of implant users to discriminate complex pitch direction required a mean threshold interval of 7.6 semitones, whereas normal-hearing listeners required only 1.1 semitones. Interestingly, this study showed that the ability to discriminate pitch correlated only moderately with melody recognition, suggesting that pitch discrimination was necessary but not sufficient for melody recognition (Gfeller et al. 2002a). In a recent study involving 114 cochlear implant subjects, Gfeller et al. (2007) observed a significant difference in the probability of correct pitch ranking when comparing the use of base frequencies of 131 and 831 Hz, suggesting that tasks of pitch perception may depend on frequency. Pitch ranking referred to the ability to identify the direction of pitch change correctly. However, these results depended on the type of cochlear implant used, specifically, long electrode versus short electrode, hybrid acoustic, and electrical stimulation implants. Melody recognition has been examined through various tasks such as those involving piano tones representing simple melodies commonly known to the general U.S. public (Gfeller et al. 2002c), as well as complex excerpts from musical genres from country to classical (Gfeller et al. 2005). Timbre is defined as the combination of qualities of a sound that distinguishes it from other sounds of the same pitch and volume such as when distinguishing between the sounds of different musical instruments. Timbre recognition has been studied by asking users to identify different musical instruments playing an identical melodic line (Gfeller et al. 2002b). Regardless of the nature of the sound stimuli, implant users have performed consistently worse than normal-hearing listeners in pitch discrimination and in both melody and timbre recognition tasks.

A number of tests have been created for the testing of music perception in cochlear implant users, including batteries developed by Gfeller et al. such as the Primary Measures in Music Audiation, Music Excerpt Recognition Test, and Iowa Music Perception and Appraisal Battery (Gfeller & Lansing 1991; Gfeller et al. 1997, 2005). Many of these tests, although providing important insights into various aspects of music perception, require hours to complete, require the aid of trained personnel, and would be difficult to administer in a typical clinical setting. Studies at various institutions use different melodies in their recognition tasks, and the same melody can be presented in varying manners, from the use of different instruments and pitch registers to the use of single melodic lines and “real-world” recordings. Many of these recordings contain rhythmic cues that may contribute to melody recognition (Monahan & Carterette 1985; Palmer & Krumhansl 1987). Galvin et al. (2007) demonstrated that cochlear implant users performed significantly better on familiar melody recognition tasks when rhythmic cues were available. A lack of standardization in the testing of music perception also precludes the ability to compare results from patients across different institutions.

As cochlear implant and signal processing technology continues to advance, the development of a standardized, clinical music perception tool would provide a method by which to quantify and gauge improvement. We have developed the University of Washington Clinical Assessment of Music Perception (CAMP) test, which is intended to serve as a reliable and valid measure of music perception for adult cochlear implant users. We specifically aimed to develop a self-administered test that could be completed in half an hour, evaluating an implant user’s ability to discriminate pitch direction, recognize melodies, and distinguish timbre. We hypothesized that music perception in part reflects cochlear implant function and is consistently measurable using our test. We also established normative data for normal-hearing adults regarding difficulty and performance of each of the test items. We hypothesized that normal-hearing subjects would perform significantly better on all aspects of the music perception test compared with cochlear implant users.

MATERIALS AND METHODS

The CAMP test comprises three subtests, including pitch direction discrimination, melody recognition, and timbre recognition. They evaluate the ability to perceive the intervallic direction of pitch pairs and to identify common melodies and the sounds of various musical instruments from closed sets, respectively.

Development of Sound Stimuli

In both the pitch and melody subtests, digitally synthesized, complex tones were used to provide both fundamental and overtone frequency cues present in real-world tones. The tones were created with identical spectral envelopes from a single synthetic piano note at middle C (262 Hz) to which uniform temporal envelopes were applied to eliminate potential temporal envelope cues. A custom peak-detection algorithm was used to abstract the fast Fourier transform components corresponding to the fundamental (F0) and harmonic frequencies. Harmonic peaks were defined as those greater than 15% of the fundamental frequency, or maximum peak, and corresponding harmonic frequencies were determined within 50 Hz around each peak. Sinusoidal waves were generated using the spectral amplitudes and phase values from the fast Fourier transform peaks. These sine waves were summed and multiplied by a uniform temporal envelope characterized by exponential onset and linear decay (Nimmons et al. 2008).

Because pitch discrimination ability has been suggested to depend on frequency (Gfeller et al. 2002b; Vandali et al. 2005; Laneau et al. 2006), three base frequencies in the octave above middle C or C4 were tested: C4 at 262 Hz, E4 at 330 Hz, and G4 at 392 Hz. Although not representative of the entire musical pitch range, testing was limited to these base frequencies based on their situation within a commonly used vocal and instrumental frequency range in Western music, regular use in our melody samples, and the desire to maintain brevity of test duration.

Our melody subtest includes 12 commonly known melodies, including “Frère Jacques,” “Happy Birthday,” “Here Comes the Bride,” “Jingle Bells,” “London Bridge,” “Mary Had a Little Lamb,” “Old MacDonald,” “Rock-a-Bye Baby,” “Row Row Row Your Boat,” “Silent Night,” “Three Blind Mice,” and “Twinkle Twinkle Little Star” (Nimmons et al. 2008). These melodies were selected from a much larger initial list of candidate melodies. Discussion among clinicians and musicians was used to narrow the list to what were considered the most commonly recognizable melodies that included primarily nursery songs. All melodies were again created in the octave surrounding and above middle C. Melodies were played in an isochronous manner, in which the melodic line was played using only consistent repetitions of eighth notes. This method of melody presentation was used to eliminate rhythm cues that have been suggested to confound melody recognition (Monahan & Carterette 1985; Palmer & Krumhansl 1987). Each melody was created using a standard tempo (quarter note = 60 beats per minute) and note duration (500 msec) and was truncated to standardize melody clip duration (8 sec). The amplitude of each note was randomly roved ±4 dB to eliminate any loudness cues. Thus, we attempted to isolate the variability of pitch sequence, or melody, as the only distinguishing element in these test items.

The timbre subtest included eight musical instruments from the four major instrument classes (strings, brass, woodwinds, and percussion) playing the identical five-note sequence of C4-A4-F4-G4-C5, composed specifically for this test. This concise melodic sequence included bidirectional intervallic changes (both stepwise movements and skips) and encompassed the octave above middle C. The instruments were selected after careful discussion of the same issues by the same clinicians and musicians involved in the melody selection. The pitched percussion instrument family was represented by the piano and nylon-stringed acoustic guitar, the stringed instruments by the violin and cello, the brass instruments by the trumpet, and the woodwind instruments by the flute, clarinet, and saxophone. The recordings were performed at a professional recording studio. Instrumentalists were instructed to play at a uniform tempo of 82 beats per minute; to use the same intensity of mezzo forte, articulation (i.e., detached), and phrasing; and to avoid the use of vibrato. The amplitude of each sample was normalized using root mean square followed by peak normalization. Synthetic musical instruments samples have been used in previous studies, but even trained musicians were unable to identify them consistently (Williams 1996). The results of this timbre subtest therefore holds “real-world” implications because live instrument sequences were recorded for the creation of the test items.

Development of Test Protocol

The three components of our music perception (pitch, melody, and timbre recognition) test were assembled into a computer-driven exercise. The pitch direction discrimination subtest used a two-alternative forced choice, 1-up 1-down adaptive testing method (Levitt 1971). A base frequency and a higher pitch were played in random order. Two buttons were presented on a computer screen, and the user was instructed to select the button corresponding to the note (first or second note heard) perceived as higher in pitch. Each correct response yielded a smaller subsequent pitch interval, and each incorrect response yielded a larger interval. A reversal was defined as an incorrect response after a correct response or vice versa. To create an accurate psychometric function, a reversal at zero was automatically added by the test algorithm when the user answered correctly at a 1 semitone interval. The initial interval presented was 12 semitones, or 1 octave, and the smallest interval tested was 1 semitone. Discrimination thresholds were based on performance on each base frequency independently, but all were tested concurrently, so that presented pitch pairs randomly used any of the three base frequencies. The subject's pitch discrimination threshold at each base frequency was calculated by the mean pitch interval in semitones using the last six of eight total reversals. After three trials at each base frequency were completed, a mean threshold was calculated for each base frequency. The final pitch direction discrimination score was calculated as the mean of all three base frequency thresholds.

In the melody recognition subtest, each melody clip was played three times in random order. Subjects were asked to identify melodies by selecting the title corresponding to the melodies presented. Similarly, in the timbre (musical instrument) recognition subtest, the previously recorded instruments were presented three times in random order, and subjects were instructed to select the labeled icon of the instrument corresponding to the timbre presented. The scores on these subtests were calculated as the percent of melodies or instruments correctly identified.

No feedback was given during the entire music test. Each subtest began with a brief training session in which subjects were able to familiarize themselves with the test items and protocol at their own pace. All testing was performed in a sound-proof booth in which subjects were seated 1 m in front of a loudspeaker that delivered the sound stimuli through a free field. Mean sound levels were calibrated to 65 dBA. A computer screen and mouse were situated just to the side of the subject, so as not to obstruct the delivery of sound from the loudspeaker. Before administration of the CAMP test, subjects rated their familiarity with the title of each melody on a five-point scale, from “not familiar” to “very familiar.”

Subject Inclusion and Exclusion

Subjects included in this study were a convenience sample of volunteer, postlingually deafened, adult cochlear implant users. Subjects were referred from the University of Washington Medical Center, as well as through a volunteer network of cochlear implant users affiliated with the Cochlear Corporation and Advanced Bionics Corporation. Subjects were excluded if they were not able to complete questionnaires in English, could not perform the visual and motor tasks required of the testing components (seeing the computer screen and pressing appropriate keys), did not have adequate cognitive abilities, or could not complete the test. Normal-hearing listeners included volunteers from the campuses of both the University of Washington and Veterans Affairs Puget Sound Medical Centers and were included upon verification of normal-hearing status, defined by audiometric thresholds better than 25 dB hearing level at octave test frequencies from 250 to 8000 Hz. Approval by the Institutional Review Board of the University of Washington was obtained before the commencement of this study.

Data Collection

At the initial visit, we collected contact and demographic information, history of hearing loss and cochlear implantation, and self-assessment of musical background using questionnaires. We also administered hearing-specific outcomes measures, including the Performance Inventory for Profound and Severe Loss (PIPSL; Owens & Raggio 1988) and the Hearing Handicap Inventory for the Elderly (HHIE; Weinstein & Ventry 1983). The PIPSL is a 74-item scale that has been shown to be valid and reliable in gauging patients' subjective assessments of their performance in a variety of communicative situations. The questions characterize hearing loss across seven different categories, labeled Understand Speech with Visual Cues, Intensity, Response to Auditory Failure, Environmental Sounds, Understand Speech with No Visual Cues, Personal, and General. A lower score is indicative of higher function. The HHIE is a validated 25-item scale that assesses the emotional and social handicap from hearing impairment. CNC monosyllabic word recognition scores in quiet and speech recognition thresholds (SRT), both in steady state noise and two-talker babble, were obtained in cochlear implant subjects. SRTs were obtained using a test developed by Won et al. (2007), in which subjects identified randomly chosen spondee words out of a closed set of 12 equally difficult spondees. Finally, the CAMP test was administered. Cochlear implantees repeated the CAMP test 1 day to several weeks later to assess test-retest reliability. Normal-hearing volunteers underwent audiograms to verify normal-hearing status and then completed the music perception test and the demographic, contact, and musical background questionnaires. Normal-hearing volunteers were not retested.

Data Analysis

Tenets of classical test theory were used to establish reliability and validity of the CAMP test. Test-retest reliability was explored by administering the test to cochlear implant subjects on two separate occasions to demonstrate stability of test scores in the absence of clinical change. The three types of validity that we sought were content, criterion, and construct. Content validity was established through discussion among music professionals and clinicians about appropriately recognizable test items for the melody and timbre subtests. Criterion validity, in which test performance is compared with a preexisting gold standard, was less relevant because this study explored the development of a new metric for which no gold standard exists. Therefore, the focus of testing was on construct validity, which is demonstrated when a scale behaves according to hypothesized relationships. We hypothesized that music perception scores would correlate moderately ($r \geq 0.35$) with CNC word recognition scores and SRTs in noise. Gfeller et al. (2007) reported moderate correlations between familiar melody recognition tasks and speech recognition in noise for both broadband noise and babble. We did not expect higher values because we suspected that different aspects of cochlear implant function contributed to speech and music perception. Furthermore, a strong correlation would suggest collinearity or redundancy of the CAMP test as a measure of implant function, because speech perception tests would be able to predict music perception ability. However, we speculated that high-performing implant users with respect to speech perception would be better able to perceive various aspects of music and vice versa. We also established concurrent validity by comparing performance between cochlear implant subjects and normal-hearing volunteers.

RESULTS

Subjects

Forty-two cochlear implant users and 10 normal-hearing listeners participated (Table 1). A variety of implant models were used (Table 2). Known etiologies of hearing loss in the cochlear implant users included hereditary and known genetic causes, noise exposure, medication, otosclerosis, radiation, head trauma, meningitis, and other infection.

Performance of Subjects on CAMP Subtests

The mean duration of the CAMP was 37 min for all cochlear implant subjects. Implant users demonstrated a mean pitch direction discrimination threshold of 3.0 semitones (SD = 2.3 semitones), with a range of performance between 1 and 8.0 semitones (Fig. 1A). The mean thresholds in semitones at each of the base frequencies were 2.9 (SD = 2.7) at 262 Hz, 3.4 (SD = 3.1) at 330 Hz, and 2.5 (SD = 2.5) at 392 Hz. The pitch data for nine consecutive subjects were lost because of data file corruption. Implant users correctly recognized a mean of 25.1% (SD = 22.2%) of the melodies presented to them, with a range of performance between 0 and 94.4% (Fig. 1B). Subjects' reported mean familiarity ratings were greater than 4.0 on a five-point scale for all melodies. Melody difficulty, or the number of times a melody was correctly identified by each subject, did not correlate with reported familiarity ratings. Implantees also recognized a mean of 45.3% (SD = 16.2%) of the musical instruments presented to them, with a range between 20.8 and 87.5% (Fig. 1C). Chance performance refers to the expected performance in percent correct if melodies and instruments were chosen completely at random. For the melody recognition subtest, chance performance was 1 of 12 or 8.3%; for the timbre subtest, 1 of 8 or 12.5%.

t-Tests revealed no significant differences in subtest results between unilateral and bilateral implant users. Performance on each subtest did not differ significantly among implant manufacturers or models based on analysis of variance testing. Age and years of implant experience did not individually yield correlation coefficients greater than 0.2 with music subtest outcomes (Table 3). However, significant correlations were observed between duration of deafness before implantation and pitch threshold (Pearson's correlation coefficient [r] = 0.39, $p < 0.03$) and between duration of musical instrument experience and melody recognition ($r = 0.34$, $p < 0.04$).

Difficulty of Individual Melody and Timbre Subtest Items

Subtest item difficulty is determined by the percentage of time a given melody or musical instrument is identified correctly by all subjects from all presentations. Each melody and instrument were presented to each subject a total of three times. Among cochlear implant users in the melody subtest, item difficulty ranged from 15.7% correct recognition for "Happy Birthday" to 48.0% for "Twinkle Twinkle" (Fig. 2A). In the timbre subtest, item difficulty ranged from 25.5% correct recognition for the flute to 63.7% for the piano (Fig. 2B). Confusion matrices demonstrated that the flute was more often confused with the cello (32.1%) than correctly identified. Such interfamily confusions by implant users, particularly mistaking the flute for the cello, have been reported (Gfeller et al. 2002c). However, most confusions by implant users in this study were intrafamily. When presented to normal-hearing listeners, all melodies were recognized more than 72% of the time and all musical instruments more than 87% of the time.

Validity and Reliability Testing of CAMP Subtests

Test-retest reliability was examined. Of the 42 cochlear implant users who were tested with the CAMP, 36 were retested. Six subjects were unable to return for retesting. Many were retested over the course of 2 consecutive days. Others were only able to return up to 3 to 4 wk after their initial music testing to coincide with other scheduled medical appointments. One subject returned after 6 mo for retesting and was not included in the test-retest analysis because absence of clinical change could not be reasonably assumed. Of the remaining 35 subjects who were retested, the mean retest interval was 5.7 days (SD = 11.1 days). The intraclass coefficients for pitch, melody, and timbre subtests were 0.85, 0.92, and 0.69, respectively, demonstrating strong test-retest reliability of the pitch and melody subtests and moderate to strong reliability of the timbre subtest (Fig. 3). The mean difference between test and retest scores on the melody subtest was 2.7% (SD = 9.5); it was 6.1% (SD = 14.0) for the timbre subtest. Normal-hearing listeners were not retested because all were able to perform very well on their first test instances.

Thirty-five of 42 implant subjects completed CNC word recognition testing. Of these results, CNC scores correlated moderately with the melody and timbre subtests of the CAMP test, whereas the pitch subtest correlation was moderate to strong. SRTs in steady state noise and two-talker babble were obtained in 39 implant subjects. Each music subtest correlated moderately with both types of background noise. Pearson's correlation coefficients are shown in Table 4.

The results from each subtest correlated moderately with each other. The Pearson's correlation coefficients were −0.57 ($p < 0.0006$) for melody and pitch (Fig. 4A), 0.68 ($p < 0.0001$) for melody and timbre (Fig. 4B), and −0.64 ($p < 0.0001$) for timbre and pitch.

Normal-hearing volunteers did uniformly well on the test. Their mean pitch direction discrimination threshold was 1.0 semitone (SD = 0.3 semitones), mean melody recognition score was 87.5% (SD = 8.3%), and mean timbre recognition score was 94.2% (SD = 4.0%; Fig. 1). Two-tailed *t*-tests revealed significant differences between these groups ($p < 0.002$).

Correlations with Patient-Reported Subjective Outcomes

Twenty-six implant subjects completed the HHIE questionnaire. Scores from both the emotional and social domains of the HHIE demonstrated no correlations with any of the CAMP subtests. Twenty-five implant subjects completed the PIPSL. Again, no correlations were seen with any CAMP subtest. However, scores from the PIPSL domain, Understand Speech with No Visual Cues, correlated moderately in a statistically significant manner with both CNC word recognition scores ($r = -0.41$, $p < 0.04$) and SRTs in steady state noise ($r = 0.50$, $p < 0.02$).

DISCUSSION

We have developed and validated a clinically practical music perception test for adult cochlear implant users. Such a test requires the ability to discriminate a wide range of ability among implant users through the use of test items of varying degrees of difficulty. The CAMP test demonstrated a broad range of scores in all three subtests. Additionally, individual test items in the melody and timbre recognition subtests also demonstrated a breadth of difficulty. The intent in the creation of the CAMP test was to provide a tool that could be administered rapidly in a clinical setting without requiring a specially trained technician. As such, a user-friendly computer interface was created with self-explanatory directions requiring only the click of a mouse. In our experience, subjects were able to follow these directions easily to complete the music test in a mean testing time of 37 min, which we consider clinically feasible. All subjects were able to complete the test in its entirety. The test-retest data suggest that test performance is stable. We have demonstrated construct validity with both existing measures of speech recognition and hypothesized relationships between domains of the CAMP. Each subtest of the CAMP correlated moderately with each other, including pitch and melody. The moderate correlation suggests that although the tests are related, they reflect different aspects of music perception. For example, subjects who scored highly on melody recognition also performed well on the pitch subtest, although the converse was not always the case. This supports similar findings in the literature, suggesting again that pitch perception is necessary but not sufficient for melody recognition.

Although correlations between CNC word recognition and our music subtests behaved according to hypothesized relationships, it is difficult to establish a similar construct against which music perception can be compared for several reasons. First, elements of music perception are distinct and diverse. The ability to identify pitch differences and various melodies and timbres through an implant most likely depends on both patient and implant signal-processing factors. Although data were collected on musical background and intentional music listening to ascertain the effects of these intangible factors on music perceptual ability, only experience with a musical instrument correlated weakly to moderately with melody recognition. An individual's musical inclination is not necessarily summarized in such pieces of information as years of instrument lessons. The nature of the innate musical ability required for melody or timbre recognition and pitch direction discrimination is not well defined and not easily quantified, but experience with a musical instrument may be beneficial.

Although we have attempted to isolate specific elements of music that we considered critical to music perception, we may not have accounted for every unintended cue that could have existed in our sound stimuli. Items in the melody recognition subtest were played in an isochronous manner to eliminate rhythmic cues; each tone was synthetically generated to ensure uniformity of temporal envelope; and each note was randomly roved ±4 dB to eliminate loudness cues. However, despite our attempts at controlling these variables and thereby creating a solely pitch-dependent melody recognition test, it is possible that other perceptual discrepancies were available to the implant user. For example, characteristic sound quality changes corresponding to pitch patterns in a given melody line may have been recognized by subjects, as opposed to the accurate perception of the pitches themselves. Similarly, our timbre recognition subtest used live instrument recordings because we were careful to present real sounds in a test whose purpose was to assess the ability to distinguish particular timbres. Therefore, it was expected that some characteristics such as onset and offset transients that define unique timbres, along with minute incidental differences in the performance of each instrument, might have provided cues to the implant user. Longer, sustained notes, while providing more information to listeners, would create opportunity for additional instrumentalist-dependent, note-to-note variability that would furnish additional cues. Although some cochlear implant subjects in our study reported being able to identify melodic lines and the characteristic sounds of musical instruments, others reported simply learning to associate various patterns in the sound stimuli with the tested melody and timbre items during the practice portion of these subtests. Whichever may be the case for a given implant user, the test remains a useful means by which to stratify cochlear implant listeners' abilities to identify such musical cues. Discrimination of timing, or rhythm, in implant users has been demonstrated as nearly normal (Shannon 1993). Rhythm is represented in temporal envelope coding in implants, which is quite good. For these reasons and to create a more efficient music test, rhythm testing has been excluded from the University of Washington CAMP.

Another limitation of our melody and timbre recognition subtests is their potential reliance on subject's prior knowledge of the included songs and instruments. These subtests might therefore pose more of a challenge to prelingually deafened implant users. However, subjects with appreciable music perception ability may potentially learn the melodies and sounds of the instruments during the practice portion of the subtests. In fact, it can be expected that even normal-hearing listeners may not be completely familiar with the characteristic sounds of the various woodwind instruments, for example. However, the ability to distinguish timbres is nevertheless tested as subjects attempt to familiarize themselves with the different instrument sounds during the practice module. To limit the uncontrollable effect of subject unfamiliarity with melodies and timbres, care was taken to select only what we considered to be highly recognizable melodies and instruments. All subjects in this study reported a mean familiarity rating between 4 and 5 on a five-point scale for all melodies except "Frere Jacques," which scored a mean rating of 3.9. However, self-reported familiarity ratings did not correlate with correct recognition of melodies; "Happy Birthday" scored the highest mean recognition rating but was the least correctly identified melody, and a subject who reported no prior familiarity with "Frere Jacques" correctly identified the melody at every presentation. None of the normal-hearing volunteers, who had cultural backgrounds similar to those of the implant subjects, had difficulty with melody recognition.

The learning effects of multiple test renditions have not been fully explored. Test-retest reliability data from this study suggest that learning is minimal. The mean retest interval was 5.7 days, with many subjects retesting over 2 consecutive days because of scheduling restrictions. This brief interval favors learning, so the test-retest reliability data are likely skewed toward improvement in performance. Despite this situation, our test-retest reliability correlations were strong.

The explanation for the range of difficulty observed in the individual melody items is not readily apparent. It might be hypothesized that those melodies with large intervallic jumps might be more easily distinguished. However, “Happy Birthday” was one of the melodies with the largest pitch ranges and intervallic jumps yet was the least recognized melody. “Mary Had a Little Lamb,” on the other hand, had one of the smallest pitch ranges and was the fourth most recognized.

The pitch direction discrimination thresholds of the cochlear implant listeners in this study were lower than those described by Gfeller et al. (2002a), with a mean of 3.0 semitones (SD = 2.3) compared with 7.6 semitones (SD = 5.2). This may be explained by differences between the two populations, testing algorithms, and range of pitches tested. The range of pitches tested in this study was 19 semitones, whereas the range in the previous study was 36 semitones.

Although other valuable tests exist that provide insight into aspects of music perception different from those explored in the CAMP test, we have created a brief, self-administered tool examining a few key elements of music perception with utility in both research and clinical environments. As speech perception continues to improve, music perception remains problematic for most cochlear implant users. With the prevalent desire for improved music appreciation in this population, the development and implementation of new implant technology and processing strategies should be examined with respect to their capacity for conveying elements of musical sounds as well as speech. Just as speech perception measures are routinely collected after implantation and adjustments in processing strategies and maps, the addition of music perception testing in outcomes assessments would also provide important insights into such changes.

The rapid nature and ease of administration of the CAMP test provide a tool for the standardized assessment of music perception outcomes in cochlear implant users. A standardized method of testing would not only allow familiarization with a particular music perception test and the interpretation of its results but also enable the comparison of outcomes in the greater cochlear implant population across multiple institutions. This would facilitate the same type of generalizable assessment made possible for speech outcomes by the establishment of the standardized minimum speech test battery including CNC words and Hearing in Noise test (Luxford 2001). In light of these implications for the uses of the CAMP test, the ability of this tool to detect clinical change should be explored through responsiveness testing in future studies. Our experience with the CAMP suggests that it may be a reliable and valid test of music perception in adult cochlear implant users. Its use affords the opportunity to establish quantifiable benchmarks by which implant technology may continue to improve the communication of music and by which clinicians may make judgments regarding the use of different processing strategies. The CAMP test offers a role in the assessment and treatment of sensorineural hearing loss because music perception continues to remain a pivotal issue in the aural rehabilitation of cochlear implant patients.

Acknowledgments

The authors thank the subjects for their efforts in completing this extensive testing protocol, as well as Jay Kenney of Audio Logic Inc. for his expertise in studio recording and sound editing.

This work was supported by NIH grants R01-DC007525, P50-DC00242, T32-DC00018, P30-DC004661, and AAO-HNSF; Cochlear Corporation; and Advanced Bionics Corporation. Dr Rubinstein is a paid consultant for Cochlear Corporation and Advanced Bionics Corporation, two manufacturers of cochlear implant systems.

References

Friesen LM, Shannon RV, Baskent D, et al. Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. J Acoust Soc Am. 2001; 110:1150–1163. [PubMed: 11519582]

Galvin JJ, Fu QJ, Nogaki G. Melodic contour identification by cochlear implant listeners. Ear Hear. 2007; 28:302–319. [PubMed: 17485980]

Gfeller K, Christ A, Knutson JF, et al. Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. J Am Acad Audiol. 2000; 11:390–406. [PubMed: 10976500]

Gfeller K, Knutson JF, Woodworth G, et al. Timbral recognition and appraisal by adult cochlear implant users and normal-hearing adults. J Am Acad Audiol. 1998; 9:1–19. [PubMed: 9493937]

Gfeller K, Lansing CR. Melodic, rhythmic, and timbral perception of adult cochlear implant users. J Speech Hear Res. 1991; 34:916–920. [PubMed: 1956198]

Gfeller K, Olszewski C, Rychener M, et al. Recognition of “real-world” musical excerpts by cochlear implant recipients and normal-hearing adults. Ear Hear. 2005; 26:237–250. [PubMed: 15937406]

Gfeller K, Turner C, Mehr M, et al. Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. Cochlear Implants Int. 2002a; 3:31–55.

Gfeller K, Turner C, Oleson J, et al. Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. Ear Hear. 2007; 28:412–423. [PubMed: 17485990]

Gfeller K, Witt S, Adamek M, et al. Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients. J Am Acad Audiol. 2002b; 13:132–145. [PubMed: 11936169]

Gfeller K, Witt S, Woodworth G, et al. Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. Ann Otol Rhinol Laryngol. 2002c; 111:349–356. [PubMed: 11991588]

Gfeller K, Woodworth G, Robin DA, et al. Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. Ear Hear. 1997; 18:252–260. [PubMed: 9201460]

Kong YY, Cruz R, Jones JA, et al. Music perception with temporal cues in acoustic and electric hearing. Ear Hear. 2004; 25:173–185. [PubMed: 15064662]

Laneau J, Wouters J, Moonen M. Improved music perception with explicit pitch coding in cochlear implants. Audiol Neurootol. 2006; 11:38–52. [PubMed: 16219993]

Leal MC, Shin YJ, Laborde ML, et al. Music perception in adult cochlear implant recipients. Acta Otolaryngol. 2003; 123:826–835. [PubMed: 14575398]

Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am. 1971; 49(Suppl 2): 467. [PubMed: 5541744]

Luxford WM. Minimum speech test battery for postlingually deafened adult cochlear implant patients. Otolaryngol Head Neck Surg. 2001; 124:125–126. [PubMed: 11226944]

Monahan CB, Carterette EC. Pitch and duration as determinants of musical space. Music Perception. 1985; 3:1–32.

Nilsson M, Soli SD, Sullivan JA. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am. 1994; 95:1085–1099. [PubMed: 8132902]

Nimmons GL, Kang RS, Drennan WR, et al. Clinical assessment of music perception in cochlear implant listeners. Otol Neurotol. 2008; 29:149–155. [PubMed: 18309572]

Owens E, Raggio M. Performance inventory for profound and severe loss (PIPSL). J Speech Hear Disord. 1988; 53:42–56. [PubMed: 3339867]

Palmer C, Krumhansl CL. Independent temporal and pitch structures in determination of musical phrases. J Exp Psychol Hum Percept Perform. 1987; 13:116–126. [PubMed: 2951485]

Peterson GE, Lehiste I. Revised CNC lists for auditory tests. J Speech Hear Disord. 1962; 27:62–70. [PubMed: 14485785]

Shannon, R. Psychophysics. In: Tyler, RS., editor. Cochlear Implants: Audiological Foundations. San Diego, CA: Singular Publishing, Inc; 1993.

Vandali AE, Sucher C, Tsang DJ, et al. Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies. J Acoust Soc Am. 2005; 117:3126–3138. [PubMed: 15957780]

Weinstein BE, Ventry IM. Audiologic correlates of hearing handicap in the elderly. J Speech Hear Res. 1983; 26:148–151. [PubMed: 6865371]

Williams DA. A study of internal validity of the instrument timbre preference test. J Res Music Educ. 1996; 44:268–277.

Wilson, BS. Cochlear implant technology. In: Niparko, JK., editor. Cochlear Implants: Principles & Practices. New York, NY: Lippincott Williams & Wilkins; 2000. p. 109-118.

Won JH, Drennan WR, Rubinstein JT. Spectral-ripple resolution and speech perception in noise by cochlear implant listeners. J Assoc Res Otolaryngol. 2007; 8:384–392. [PubMed: 17587137]

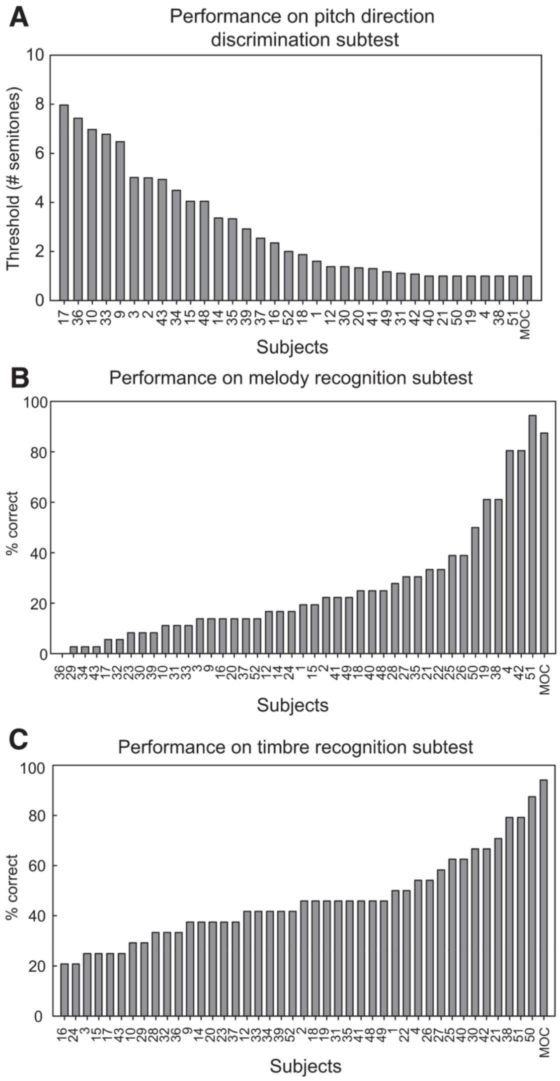

Fig. 1.

Distribution of performance on CAMP subtests (MOC, mean of controls): (A) pitch direction discrimination, (B) melody recognition, and (C) timbre recognition.

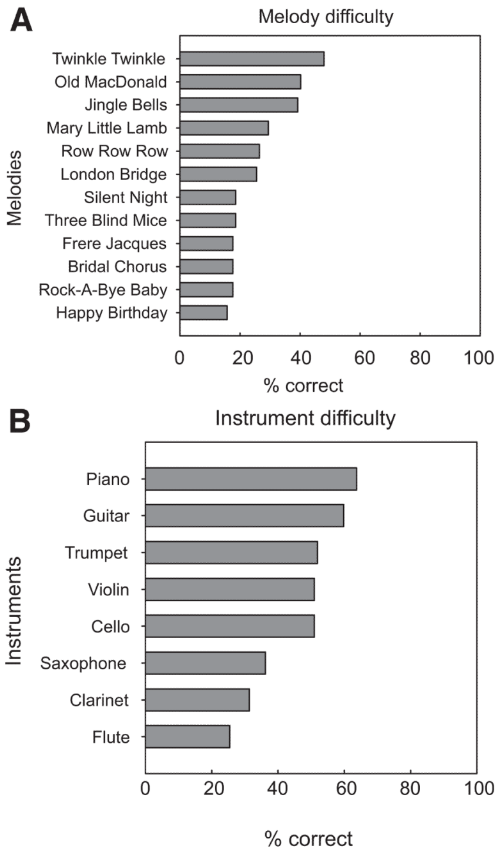

Fig. 2.

Subtest item difficulty for cochlear implant subjects: (A) melody recognition and (B) timbre recognition.

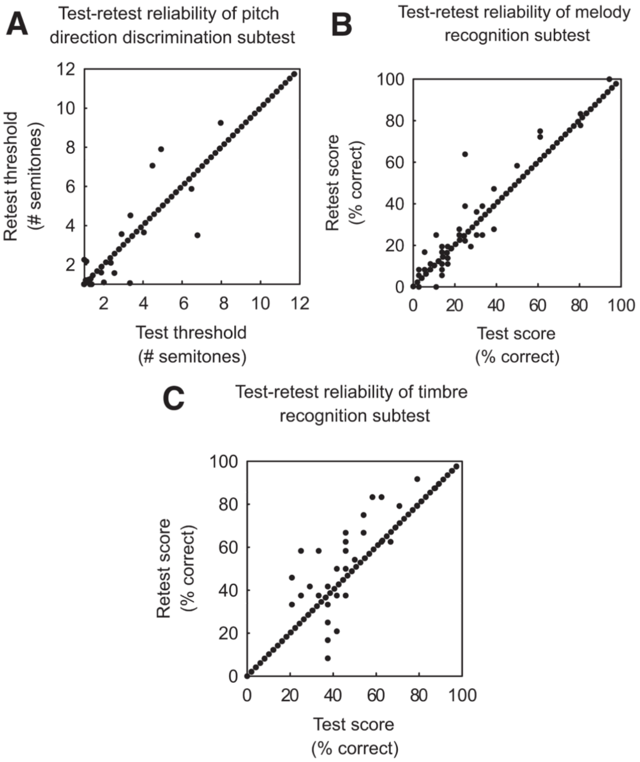

Fig. 3.

Test-retest reliability of each subtest: (A) pitch direction discrimination, (B) melody recognition, and (C) timbre recognition.

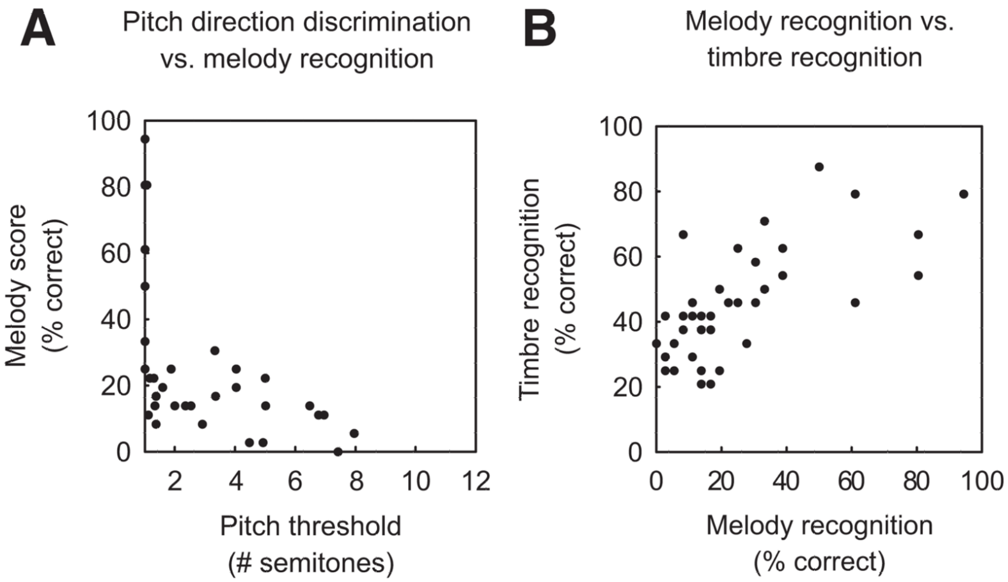

Fig. 4.

Relationships between subtests: (A) pitch direction discrimination versus melody recognition and (B) melody recognition versus timbre recognition.

TABLE 1

Subject characteristics

	Implant users	Normal-hearing listeners
N	42	10
No. of males	21 (50%)	4
Mean age (range)	58 yr (35–81)	39 (26–62)
No. of bilateral implantees	6 (14%)	N/A
Mean duration of deafness (range)	9 yr (0–57)	N/A
Mean implant experience (range)	4 yr (0–15)	N/A

TABLE 2

Frequency of various cochlear implants used by implant subjects

Implant model	Frequency
CI24R(CS)	15
CI22	8
CI24M	7
Nucleus Freedom	4
HiRes 90k	9
Clarion CII	4
Med-EL Combi 40+	1
Total	48*

*Six implantees wore bilateral implants.

TABLE 3

Pearson’s correlation coefficients showing relationships between patient characteristics and subtest results

	Melody	Timbre	Pitch
Age	0.047	0.040	−0.074
Duration of deafness	−0.19	−0.13	0.39*
Implant experience	−0.0064	0.14	0.20
Musical instrument experience	0.34*	0.17	−0.080

*Significant at $p < 0.05$.

TABLE 4

Pearson's correlation coefficients for individual subtests and CNC and speech reception thresholds in both steady state noise and two-talker babble

Subtest	CNC	Steady state noise	Two-talker babble
Pitch	-0.66^{*}	$0.55^{\dagger}$	0.58^{*}
Melody	$0.47^{\dagger}$	$-0.46^{\dagger}$	$-0.42^{\dagger}$
Timbre	$0.50^{\dagger}$	-0.53^{*}	$-0.47^{\dagger}$

[*] Significant at $p < 0.001$.

[†] Significant at $p < 0.01$.